DOCUMENT RESUME

ED 071 759 PS 006 303

AUTHOR	Gladkowski, Gerald James
TITLE	Another Look at Compensatory Education.
PUB DATE	[71]
NOTE	15p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Cognitive Processes; *Compensatory Education Programs; Disadvantaged Youth; *Evaluation Criteria; *Evaluation Methods; Experimental Programs; Literature Reviews; Measurement Instruments; Preschool Education; Preschool Evaluation; Preschool Learning; Preschool Programs; *Program Evaluation; Reliability; Research Criteria; *Research Methodology
IDENTIFIERS	Elementary Secondary Education Act Title I; ESEA Title I; Project Head Start

ABSTRACT

In considering the apparent failure of the two major compensatory education programs, Project Head Start and Title I, there are so many uncontrolled variables interacting simultaneously that it is virtually impossible to define the specific etiological factors engendering the results obtained. The Federal programs are difficult to evaluate because their goals are broad, they involve millions of children, and they are administered by Federal agencies far away from individual projects. Other problems are created through such variables as program effects or maturation (our lack of knowledge about preschool learning and disadvantaged learners), interactions of various socializing agencies, and technology. Reliability of measurement devices is especially doubtful at the preschool level. Our lack of knowledge in the affective domain is even greater. The major weaknesses of the compensatory evaluation are: (1) lack of comparable groups and control groups; (2) no planned variation in programs; (3) lack of random selection and/or assignment of Ss to treatment and control groups; (4) lack of clear-cut criteria for inclusion in the program; (5) lack of clearly specified objectives; and (6) non-comparable data. Future intervention programs should adhere to the tenets of experimental research. Residential centers are recommended to remove disadvantaged children from impoverished environments in infancy. A bibliography is provided. (KM)

FILMED FROM BEST AVAILABLE COPY

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED 071759

Another Look at Compensatory Education

by

Gerald James Gladkowski

PS 006303

Introduction: Arthur Jensen (1969, p. 1) stated that, "Compensatory education has been tried and apparently it has failed." This statement is accurate if restricted to the overall results of the two major federally funded compensatory education programs (Project Head Start and Title I); however, it is a rather gross oversimplification of the problem. In a sense it is analagous to a research problem in that the results are not (or should not be) an end point, but, instead, a beginning or impetus to further investigation regarding "why" the effort "failed" or "succeeded." (See Edwards, 1968) Jensen (1969) contended that the failure was due to reasons extrinsic to the programs themselves, whereas the present author claims that there are so many variables interacting simultaneously that it is virtually impossible to delineate the specific etiological factors engendering the results obtained. The author further contends that the "apparent failure" could be attributed to two other equally as plausible factors, namely: (1) the actual implementation and design of the specific programs, or non-evaluational factors,e.g.,type of program, and (2) the experimental design or evaluational considerations,e.g., non-comparable groups. It should be noted that neither of these factors is independent or mutually exclusive of the other.

Because it is readily acknowledged that the results of the two major compensatory efforts have not been successful in their attempts to reach the program objectives (See Mc Dill, et al., 1969; Cohen, 1970; and Westinghouse-Ohio Study, 1969), the author will not analyze the specific findings in detail. The important point is not that these programs have "failed" but, rather, the many problems they have confronted (ranging from the political to the non-comparability of groups) and which have thus served as competing alternative hypotheses. Further discussion of the general and specific problems confronting these programs will illustrate the author's contention regarding the difficulty in delineating the many possible etiological factors contributing to the results obtained.

General Factors: Cohen (1970) indicated that prior to 1964 educational evaluation had been primarily confined to small scale research in which the purpose of the study was generally limited to rather specific factors and typically involved a small budget and staff. However, after 1964 the federal government became involved in establishing broad educational programs which Cohen (1970, p. 213) perceived as differing from the previously conducted research in three important ways:

(1) they are social action programs, and as such are not focused narrowly on teachers' in-service training or on a science curriculum, but aim broadly at improving education for the disadvantaged.

(2) the new programs are directed not at a school or a school district, but at millions of children, in thousands of schools in hundreds of school jurisdictions in all the states;

(3) they are not concerned and executed by a teacher, principal, a superintendent, or a researcher - they were created by the Congress and are administered by federal agencies far from the school districts which actually design and conduct the individual projects.

Without delineating all the questions and implications engendered by the above it is obvious that any large scale program will create many problems. For example, how does one effectively evaluate the specific effects upon approximately three million children spread out across the nation? Is it reasonable to evaluate on the basis of criteria related primarily to achievement for programs directed at rather broad, sweeping political, economic, and social changes? Should evaluation be decentralized despite the fact that they are national programs? How does one determine the specific effects of any undertaking when the overall objectives for the program are determined nationally, but yet each local school district (or state agency) is responsible for the actual implementation of the program? These are but a few of the important questions that could be raised and as Cohen (1970, p. 215) stated, "In the social action programs, however, the political importance of information is raised to a high level by the broader political character of the programs themselves." The important point being that while the basic tenets of experimental research may be similar for evaluating both small and large scale programs,i.e., determining their effects, the important difference lies in the character of the aims and organization of the programs. Campbell (1969, p. 410) reached conclusions similar to Cohens, stating that, "If the political and administrative system has committed itself in advance to the correctness and efficacy

of its reforms, it cannot tolerate learning of failure. To be truly scientific we must be able to experiment. We must be able to advocate without that excess of commitment that blinds us to reality testing." For example, one would logically assume that some type of evaluational procedure would be involved in order to assess whether or not a program has been effective, but as Cohen (1970, p. 219) states:

> The mandate for evaluation - like many Congressional authorizations - lacked any enabling mechanism: responsibility for carrying out the evaluation was specifically delegated to the state and local education authoritites who operated the programs. It was not hard to see, in 1965, that this was equivalent to abandoning much hope of useful program evaluation.

Campbell (1969) indicates that many feel we are at the point of continuing or discontinuing programs on the basis of assessed effectiveness. He questions the validity of this attitude when he states that, "most ameliorative programs end up with no interpretable evaluation." Another example is the fact that Title I programs are funded on a formula grant type basis, in which the amount of money given to any educational district is based on how many poor children in the district has enrolled in the schools, and not on how well the district may or may not educate. The actual implementation and evaluation of these programs are thus confounded by many non-evaluational considerations; for example, politically vested interests on various levels and the emotionally laden overtones of such programs. For a more detailed and complete discussion of other factors one is directed to Campell, 1969; Timpane, 1970; Cohen, 1970; Mc Dill, et al., 1969.

Another problem confronting compensatory education programs, specifically at the preschool level, has to do with the type of variables with which an investigator must cope. Mc Dill, et al. (1969, p. 7) cites three important variables or factors which affect compensatory programs: namely, program effects or maturation, interactions of various socializing agencies, and technology.

Many programs directed at preschool and elementary school children, are based on a common belief among some educators that the earlier we begin assisting children of this age the more successful we may be. (For example, see Bloom, 1964) The problem this creates is that we have accumulated much more knowledge of the learning process and the effects of other variables upon children in the elementary school, relatively speaking, than those that affect preschool children. Only in recent years have efforts been made to study this much younger population. According to Mc Dill, et al. (1969, p. 7), "Compensatory education or no compensatory education, we simply do not know much about how preschool children learn, and we know even less about disadvantaged learners." Because of this it is difficult to determine whether the programs themselves are ineffective, or whether they are ineffective because of our inability to define the critical variables in order to assess the impact of the program. Campbell and Stanley (1966) discuss a related problem when they list maturation as a potential confounding variables which might possibly effect the internal validity of an experiment. They ask the question, "How does one distinguish between maturation and treatment effects in young children?" Research strategies relating to the intervention model proposed by many will naturally be affected by

PS 006303

a lack of scientific knowledge concerning intellectual and general cognitive processes. Compensatory education as a strategy is not in question but rather the theorectical structure which supports the decisions that implement such a program. The present state of knowledge and the problems it creates for those interested in assessing the impact of various programs remains an obstacle to certainty in assessment. Generally speaking, researchers attempt to select one point in time as the input and another as the output, but research does not indicate if the two points are necessarily the most important in the life cycle of the individual. It may be that the significant factors have occurred prior to the experimental treatment (a problem, by the way, in all research), and is one reason why many recommend that program implementation be begun in infancy. (For example, see Bloom, 1964; Hechinger, 1966; and Gladkowski, 1971) The important point is that we do not actually know whether our programs have the effect they are designed to have. As Zimilies (1969, p. 179) stated:

> The problem, then, is reduced to finding the appropriate inputs for achieving the desired output. While schematically this may appear to be an accurate analysis of the problem, it by passes the critical intervening and mediating factor - the child. No where does one find a description of the four-year-old child, a developmental analysis of the personality and cognitive functioning of children at this age level, or a statement of their primary areas of conflict, typical modes of resolution, and principle spheres of development.

Interaction between socializing agencies represents another source of difficulty for evaluation. This problem evolves around the fact that education does not take place exclusively in the schools. A child may be involved in a formal educational program for six hours per day with the other eighteen hours representing no controlled educational activity from schools. Does the remainder of the time outside the program cancel

any potentially positive effects that might have occurred during the treatment? Is there an optimum amount of time spent in school which could be effective? The answers to these questions are, of course, not available at the present time, although they are questions which will eventually need answers if we are to identify and assess the effectiveness of our programs.

Gordon (1970) presented an excellent overview of various attempts to assist disadvantaged segments of our society in which he provides a brief synopsis concerning the areas of concern and directions for approaching the problems in program implementation for the disadvantaged. Much of the difficulty of explanation and interpretation of the various positions arise due to a confounding of factors in an attempt at delineation. For example, it has been shown that as Southern Blacks move North their achievement levels increase. The question arises, however, as to whether this is due to the impact of the school, selective migration, non-school environmental conditions, the interaction of these factors, or others not yet investigated. The interaction of many factors increases the complexity of attempts at explaining any outcome of an intervention effort. (For example, see Grotberg, 1969)

According to Mc Dill, et al. (1969), if one had a firm idea of the relevant variables important to any program design one would still be faced with the question of measurement. How much can we rely on our measurement devices to give us the data we need for evaluating outcomes? The difficulty arises at all levels but even more so at the preschool level because of the relative lack of measurement data concerning this age group. The younger the child the more inaccurate our measurment

devices are likely to be. For example, if a child were tested at age two on one of the standardized infant scales available we would not expect as high a correlation with later achievement as we would if we were to administer the test at later ages. Mc Dill, et al. (1969) indicates that while the state of development regarding cognitive dimensions is still "primitive" the picture is even more depressing when one considers the affective domain. (See Wick and Beggs, 1971; Cronbach, 1960; and Mehrens and Lehmann, 1970 for a more complete discussion of the many factors involved in the evaluational process)

The discussion presented above concerned itself with general factors affecting evaluational research. This section will delineate some of the more specific research problems (e.g., lack of control groups) relating to compensatory education programs.

One of the primary difficulties inherent in compensatory programs has been an obvious lack of control over relevant variables, ranging from non-comparable groups for comparisons, (no control groups in certain instances), to the interaction effects of the environment. (Mc Dill, et al., 1969) For example, the evaluation of Project Head Start contained many factors which were uncontrolled in the design. First, randomly selected experimental or control groups were not used but instead an ex-post-facto design in which the controls were selected and matched after the experimentals had already received the treatment constituted the basis for the evaluation. This, of course, makes it impossible to determine the specific effects of the program and thus violates one of the basic tents of experimental research. It should be indicated that the evaluators of Project Head Start did randomly select the centers

for the study but this was invalidated by the previously cited weaknesses inherent in the assessment of various local programs. The following factors were cited as representative of these weaknesses: (Westinghouse-Ohio Study, 1969)

1. Lack of comparability among separate and independent studies because of different enrollment criteria, program treatments, design, instrumentation, and schedules for gathering data.
2. In some cases the absence of any comparison group.
3. Too few cases, frequently only those enrolled at a particular center.
4. Geographical restrictions to local or regional groups.

On the basis of these difficulties selecting a "random" sample of an already biased or non-comparable sample does not eliminate the sources of bias.

Second, there were no uniform or standardized procedures adhered to between various programs to insure that the evaluation would be attempting to assess those factors which programs shared in common. For example, the various centers employed somewhat different goals, treatments, and program procedures, thus masking between and within center differences. Some centers were in operation for two hours per day whereas others were in operation for four hours; some centers were only in operation for two months whereas others were in operation for eight or nine months out of the year. (See Cohen, 1970 and Mc Dill, et al., 1969) Despite these differences the programs were all evaluated <u>as if</u> they were similar; however, there is no way of ascertaining which specific centers were relatively successful as compared to those which were not. Regarding this masking effect Cohen (1970, p. 226) stated that, "The problem, then, is not only to identify what the programs

deliever, but also to systematically experiment with strategies for affecting school outcomes...The movement toward experimentation presumes that the most efficient way to proceed is systematic trial and discard, discovering and repeating effective strategies." Others who hold similar views regarding "planned variations" include Smith and Light (1970) and Campbell (1969). This approach was not employed in the Head Start Project although the evaluative team did recommend this for future consideration.

In the assessment of Project Head Start the emphasis was on "overall" effectiveness of the program, disregarding those centers which might have been particularly effective. What this would mean in practice is that if a center (or certain aspects of a center) were found to be particularly effective then one could further investigate it in order to determine how it differs from the other centers or programs in its operation. If significant differences were detected then other centers could be organized in which the best features of proven programs could be incorporated, as well as the fact that presently operating programs could thus be modified.

Other weaknesses which contributed to the overall evaluational efforts included lack of uniformity across the various centers regarding such matters as the use of the same indices of measurement, objectives of the program, and the selection criteria of Ss for treatment and control groups. This uniformity had not been accomplished in many of the programs because, in part, the local programs were permitted the freedom to not only evaluate their own programs but also to decide upon a specific implementation course. As stated by Cohen (1970, p. 227), "The Office of Education... does not require that the same tests be used in all Title I projects,

indeed, it does not require that any tests be used." In order for an appropriate evaluation to be undertaken such matters as this must be considered before the implementation of the program, thus obviating later problems arising regarding interpretation of the results.

Many of the weaknesses inherent in the experimental designs are those related to internal validity, that is, those factors associated with the question: Did the experimental treatments make a difference in this specific experimental instance? (See Campbell and Stanley, 1966) With so many weaknesses in evidence it is virtually impossible to answer this question. Hence the studies undertaken to date are of very limited scientific value in determining whether or not the programs ere effective. The following comprises the major weaknesses of compensatory evaluations and would thus serve as a rather formidable list of competing alternative hypotheses to any research undertaking:

1. Lack of comparable groups, and in some cases, no control groups at all.
2. No planned variation in programs in order to assess both within and between center differences.
3. Lack of random selection and/or assignment of Ss to treatment and control groups.
4. Lack of clear cut criteria for inclusion into the program.
5. Lack of clearly specified objectives.
6. Non-comparable data,i.e., different indices of measurement.

Summary: The myriad of problems have not proven amenable to solutions with the delineation of the specific factors contributing to the problem being an almost impossible task due to the many uncontrolled variables interacting simultaneously, thereby confoudning both process and expected results. The attempts at assessment have consistently violated the sine qua nons of experimental research, with the evidence regarding the effectivness of our programs being ambiguous. (For example, see Cohen, 1970; Mc Dill, et al., 1969; Campbell, 1969; Posner, 1968; Etzioni, 1968; Hyman and Wright, 1967). In response to the initial question posed the author thus agrees with Jensen (1969) that the programs leave a great deal to be desired, although he does not agree with his premise. The fact is that the specific etiological factors remain somewhat problematic.

Alternative Strategy: Given this state of affairs the author is recommending an alternative strategy designed to improve upon two very important dimensions of any compensatory program, namely, the evaluational and programmatic.

First, an approach advocated by Campbell (1969) and Campbell and Erlebacher (1970) in which they recommend that future intervention programs adhere to the basic underlying tenets of experimental research and which closely approximate "true" experimental designs,e.g., random selection and assignment of Ss to treatment and control groups. (See Campbell and Stanley, 1966 for a more complete discussion of what is meant by "true" experimental designs and the underlying tenets of experimental research)

Second, implementing residential centers which would employ a total human development approach of assistance for children and families from impoverished milieus by removing the child in infancy (or a very early age) thereby: [1].

1. Providing an opportunity to assist a child in an environment in which he is relatively free of the pressures and tensions which exist (or might) in his usual social and family life, and which make difficult (or impossible) an evaluation of all the factors that contribute to his problems.

2. Providing an opportunity for the parents of the child to receive the needed assistance,e.g., vocational and/or educational training, while care is being provided for the children.

The author believes that the strategy recommended would improve considerably upon the evaluational and programmatic dimensions of an intervention program when compared with our previous efforts. The implications of employing a residential approach evolve around the concept of "control." That is, by placing children in such centers it would be possible to exert considerably more control over the various environmental contingencies and would thus minimize the heretofore uncontrolled sources of variance. Second, it would be a program which would assist both child and parent simultaneously thereby meeting the needs of all parties concerned.

1. For a complete discussion of this strategy see the author's thesis entitled "Residential Human Development Centers to Assist Children and Families from Deprived Milieus," submitted to the faculty of Northwestern University, June, 1971.

Bibliography

Bloom, B. S. Stability and Change in Human Characteristics. New York: John Wiley and Sons, Inc., 1964.

Campbell, D. T. "Reform as Experiments," American Psychologist, April, 1969, Volume 24, Number 4.

Campbell, D. T., and Erlebacher, A. "How Regression Artifacts in Quasi-Experimental Evaluations Can Mistakenly Make Compensatory Education Look Harmful," in Hellmuth, J. (ed.), Compensatory Education: A National Debate, Volume III of the Disadvantaged Child. New York: Brunner-Mazel, 1970.

Campbell, D. T., and Stanley, J. C. Experimental and Quasi-Experimental Designs for Research. Second Printing. Chicago: Rand McNally and Company, 1966.

Cohen, D. K. "Politics and Research: Evaluation of Social Action Programs in Education," Review of Educational Research, April, 1970, Volume 40, Number 2, 213-258.

Cronbach, L. J. Essentials of Psychological Testing. New York: Harper and Row, 1960.

Edwards, A. L. Experimental Design in Psychological Research. Third Edition. New York: Holt, Rinehart and Winston, 1968.

Etzioni, A. "Shortcuts to Social Change?", The Public Interest, 1968, 12, 40-51.

Gladkowski, G. J. "Residential Human Development Centers to Assist Children and Families from Deprived Milieus." (Doctoral thesis submitted to the faculty at Northwestern University, Evanston, Illinois, June, 1971)

Gordon, E. W. "Introduction," Review of Educational Research, February, 1970, Volume 40, Number 1.

Grotberg, E. Critical Issues in Research Related to Disadvantaged Children. Princeton, New Jersey: Educational Testing Service, 1969.

Hechinger, F. M. Preschool Education Today. New York: Doubleday and Company, 1966.

Hyman, H. H., and Wright, C. R. "Evaluating Social Action Programs," in Lazarsfeld, P. H., et al. (eds.), The Uses of Sociology. New York: Basic Books, 1967.

Jensen, A. "How Much Can We Boost IQ and Scholastic Achievement?," Harvard Educational Review, Reprint Series No. 2, October, 1969.

Mc Dill, E. L., Mc Dill, M. S., and Sprehe, J. T. Strategies for Sucess in Compensatory Education: An Appraisal of Evaluation Research. Baltimore, Maryland: The Johns Hopkins Press, 1969.

Mehrens, W. A., and Lehmann, I. J. Standardized Tests in Education. New York: Holt, Rinehart and Winston, Inc., 1969.

Posner, J. Evaluations of "Successful" Projects in Compensatory Education. U. S. Office of Education, Office of Planning and Evaluation. An occasional paper, Number 8, April, 1968.

Smith, P. V., and Light, R. J. "Choosing a Future: Strategies for Designing and Evaluating New Programs," Harvard Educational Review, February, 1970, Volume 40. Number 1.

Timpane, P. M. "Educational Experimentation in National Social Policy," Harvard Educational Review, Novemeber, 1970, Volume 40, Number 4.

Westinghouse Learning Corporation-Ohio University. "The Impact of Headstart: An evaluation of the Effects of Head Start on Children's Cognitive and Affective Development," ERIC Document Resume, ED 036 321, June, 1969. (Executive Summary)

Wick, J., and Beggs, D. L. Evaluation for Decision Making in the Schools. Boston: Houghton Mifflin Company, 1971.

Zimilies, H. "Review of Martin Deutsch's Book, The Disadvantaged Child," Harvard Educational Review, 1969, 39, 177-180.